

日本国特許庁

PATENT OFFICE
JAPANESE GOVERNMENT

JC135 U.S. PTO
09/551273
04/18/00

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出願年月日
Date of Application:

1999年 4月28日

出願番号
Application Number:

平成11年特許願第121518号

出願人
Applicant(s):

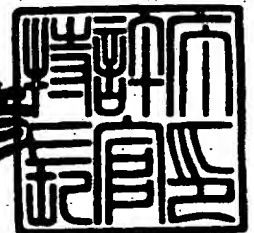
アルパイン株式会社

CERTIFIED COPY OF
PRIORITY DOCUMENT

1999年12月10日

特許庁長官
Commissioner,
Patent Office

近藤隆彦



【書類名】 特許願

【整理番号】 IWP98141

【提出日】 平成11年 4月28日

【あて先】 特許庁長官殿

【国際特許分類】 G10L 3/02

【発明の名称】 マイクロホンシステム

【請求項の数】 2

【発明者】

 【住所又は居所】 東京都品川区西五反田1丁目1番8号 アルパイン株式会社内

 【氏名】 斉藤 望

【発明者】

 【住所又は居所】 東京都品川区西五反田1丁目1番8号 アルパイン株式会社内

 【氏名】 中田 孝一

【発明者】

 【住所又は居所】 東京都品川区西五反田1丁目1番8号 アルパイン株式会社内

 【氏名】 木内 真吾

【特許出願人】

 【識別番号】 000101732

 【氏名又は名称】 アルパイン株式会社

【代理人】

 【識別番号】 100084711

 【弁理士】

 【氏名又は名称】 齋藤 千幹

 【電話番号】 043-271-8176

【手数料の表示】

 【予納台帳番号】 015222

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 マイクロホンシステム

【特許請求の範囲】

【請求項 1】 2 つマイクロホンの出力信号を用いて適応信号処理を行い SN 比を改善した話者音声信号を出力するマイクロホンシステムにおいて、

各マイクロホンを接近して配置すると共に、一方のマイクロホンから出力する信号の SN 比を高くし、他方のマイクロホンから出力する信号の SN 比を低くすることを特徴とするマイクロホンシステム。

【請求項 2】 前記一方のマイクロホンを話者の顔の真上に配置し、他方のマイクロホンを該一方のマイクロホンの位置より 1 ～ 5 cm 程度後頭部側に離して配置することを特徴とする請求項 1 記載のマイクロホンシステム。

【発明の詳細な説明】

【 0 0 0 1 】

【発明の属する技術分野】

本発明はマイクロホンシステムに係わり、特に、2 つマイクロホンの出力信号を用いて適応信号処理を行い SN 比を改善した話者音声信号を出力するマイクロホンシステムに関する。

【 0 0 0 2 】

【従来の技術】

現在の音声認識システムは、15dB 以上の SN 比 (S : 音声 / N : ノイズ) が確保されている場合、約 95 % の認識率を実現できるくらいの技術レベルにまで達している。しかし、周囲に存在するノイズにより SN 比が低下すると、それに伴って認識率が急激に低下する性質も有している。図 8 は SN 比と認識性能との関係をいくつかの種類のマイクロホン (無指向性、単一指向性、狭指向性、AMNOR (Adaptive Microphone-array for Noise Reduction)) について評価したもので、SN 比と認識率がおおむね S 字特性 100 を示す帯の中に包含されている。この図 8 から明らかなように、認識率は SN 比の低下により急激に低下し、SN 比が 0 dB の環境下において約 50 % にまで低下してしまう。

【 0 0 0 3 】

そのため、自動車が発生するノイズ（エンジン音・ロードノイズ・パターンノイズ・風切り音など）が存在する自動車車室内において、上記のような認識性能の劣化は避けられず、音声認識システムを車載化する上で大きな問題の一つとなっている。

前記したような事情から、周囲に存在するノイズの影響を少なくし、高いSN比で音声を受音するための方式が種々提案されており、複数のマイクロホンとデジタル信号処理を用いた高SN比受音システムはその一例である。かかる高SN比受音システムの中で最も簡単な構成のものは図9に示すように2つのマイクロホンを使用するシステムであるが、他にも、Griffith-Jim型アレイやAMNORといった、より高度なシステムが提案されている。

【0004】

図9において、1、2は第1、第2のマイクロホン、3は適応信号処理部であり、誤差信号 e が入力されると共にマイクロホン2の出力信号 x_2 が参照信号として入力され、誤差信号 e のパワーが最小となるようにLMS(Least Mean Square)アルゴリズムに基づいて適応信号処理を行う。適応信号処理部3において、3aはLMS演算部、3bは例えばFIR型デジタルフィルタ構成の適応フィルタである。LMS演算部3aは適応信号処理により誤差信号 e のパワーが最小となるように適応フィルタ3bの係数を決定する。

【0005】

4はマイクロホン1から出力されるノイズ信号を目標信号として入力される目標応答設定部であり、音響系の逆特性を精度よく近似するためのものである。適応フィルタ3bのタップ長の半分の信号遅延時間を d とすると、目標応答設定部4は該時間 d の遅延特性を有し、オーディオ周波数帯域でフラットな特性（ゲイン1の特性）を有する。すなわち、目標応答設定部4は、図10（a）に示すようにゲイン1のフラットな周波数特性を備え、図10（b）に示すように遅延時間 d を有するインパルス応答特性を有している。この目標応答設定部4は、FIR型デジタルフィルタの遅延時間 d に対応する係数を1にし、他の係数を0にすることにより実現できる。

5は減算部であり、目標応答設定部4から出力する目標応答より適応フィルタ

3 b の出力信号を減算して誤差信号 e を出力する。

【0 0 0 6】

非音声認識時、マイクロホン 1、2 にはノイズのみが入力し、適応信号処理部 3 は適応信号処理により誤差信号 e 、すなわち、ノイズ出力のパワーが最小となるようにフィルタ係数 W を決定する。一方、音声認識時には、適応信号処理部 3 はフィルタ係数の更新をせず、前記非音声認識時に決定したフィルタ係数 W を適応フィルタ 3 b に設定して音声信号を出力する。

図 9 に示すシステムに本来求められている理想的な性能は、音声認識時に出力信号として音声信号 $X_s(z)$ のみ(ノイズ出力は 0)を出力することである。すなわち、ノイズ出力 $E_n(z)$ に関して、

$$E_n(z) = X_{n_1}(z)z^{-d} - X_{n_2}(z)W(z) \quad (1)$$

$\{E_n(z)\}^2$ の平均が最小値をとるように調整可能なパラメータ(適応フィルタ 3 b の係数) W を決定することである。ただし、 $X_{n_1}(z)$ 、 $X_{n_2}(z)$ はマイクロホン 1、2 の出力信号に含まれるノイズ信号である。

【0 0 0 7】

【発明が解決しようとする課題】

自動車には騒音源が多数存在するため、マイクロホン 1、2 が拾う自動車車室内ノイズのコヒーレンスは、マイクロホン 1、2 を遠ざけるにしたがって低下する傾向を有している。このため、2 つのマイクロホン 1、2 を遠ざける程、(1) 式の条件が満たされにくくなってしまう問題が生じ、マイクロホン 1、2 はできるだけ近い位置に配置する必要がある。

ところが、2 つのマイクロホン 1、2 をできるだけ近い位置に配置すると、2 つのマイクロホンにほぼ同様の音声とノイズがそれぞれ入射する可能性が高くなり、(1) 式を満たすように適応フィルタ係数 W を決定してノイズを消去すると、音声までもが消去されてしまう。一方、音声が歪まないように適応フィルタ係数 W を決定すると、ノイズがほとんど消えず、SN 比もほとんど改善されなくなってしまうという問題が発生する。

【0 0 0 8】

かかる問題は、図 9 に示したシステム特有のものではなく、Griffith-Jim 型ア

レイやAMNORといった、より高度な高SN比受音システムを採用した場合でも、ほぼ同様に発生する。

以上から本発明の目的は、マイクロホンを2つ使用する図9のマイクロホンシステムにおいて、音声信号のSN比を改善できるようにすることである。

【0009】

【課題を解決するための手段】

上記課題は本発明によれば、2つマイクロホンの出力信号を用いて適応信号処理を行いSN比を改善した話者音声信号を出力するマイクロホンシステムにおいて、各マイクロホンを接近して配置すると共に、一方のマイクロホンから出力する信号のSN比を高くし、他方のマイクロホンから出力する信号のSN比を低くすることにより達成される。

このようにすれば、2つのマイクロホンの出力信号に含まれるノイズ $X_{n1}(z)$, $X_{n2}(z)$ を略等しくでき、一方、2つのマイクロホンの出力信号に含まれる音声信号 $X_{s1}(z)$, $X_{s2}(z)$ を異ならせることができる。従って、ノイズ信号入力時の $E_n(z)$ の2乗平均値が最小となるように適応フィルタ係数 W を決定しても、(2)式の音声出力 $E_s(z)$ は零とならず、音声信号のSN比を改善することができる。

【0010】

マイクロホンの具体的な配置例としては、一方のマイクロホンを話者の顔の真上に配置し、他方のマイクロホンを該一方のマイクロホンの位置より1～5cm程度後頭部側に離して配置する。このようにすれば、人間の音声放射特性により、マイクロホンが比較的近距離に配置されているにも拘らず、1つのマイクロホンはできるだけ高いSN比で音声を拾い、もう一方のマイクロホンでは、できるだけ低いSN比で音声を拾うようにできる。

【0011】

【発明の実施の形態】

(a) マイクロホンシステムの構成

図1は本発明のマイクロホンシステムの構成図であり、図9のシステムと同一部分には同一符号を付している。図中、10は話者であり、例えば自動車の運転手、11、12は第1、第2のマイクロホンである。第1のマイクロホン11は

話者の顔の真上天井に配置し、第 2 のマイクロホン 1 2 は第 1 のマイクロホン位置より 1 ～ 5 cm 程度後頭部側の天井に配置する。

3 は適応信号処理部で、誤差信号 e が入力されると共にマイクロホン 2 の出力信号 x_2 が参照信号として入力され、誤差信号 e のパワーが最小となるように LMS アルゴリズムに基づいて適応信号処理を行う。適応信号処理部 3 において、3 a は LMS 演算部、3 b は FIR 型デジタルフィルタ構成の適応フィルタである。LMS 演算部 3 a は適応信号処理により誤差信号 e のパワーが最小となるように適応フィルタ 3 b の係数を決定する。適応信号処理部 3 は、非音声認識時においてのみ適応信号処理により適応フィルタ 3 b のフィルタ係数 W を決定し、音声認識時にはフィルタ係数の更新をせず、前記非音声認識時に決定したフィルタ係数 W を適応フィルタ 3 b に設定する。

4 はマイクロホン 1 1 から出力する信号を目標信号として入力される目標応答設定部で、時間 d の遅延特性を有し、かつ、オーディオ周波数帯域でフラットな特性（ゲイン 1 の特性）を有している。5 は減算部で、目標応答設定部 4 から出力する目標応答より適応フィルタ 3 b の出力信号を減算して誤差信号 e を出力する。

【 0 0 1 2 】

（b）人間の音声放射特性

図 2 は人間の音声放射特性であり、（a）は話者 1 0 の口元を含む水平面において口元から所定距離の位置における音声レベルを周波数毎に示す放射特性図、

（b）は話者 1 0 の口元を含む垂直面において口元から所定距離の位置における音声レベルを周波数毎に示す放射特性図である。図中、A は 125Hz ～ 250Hz、B は 500Hz ～ 700Hz、C は 1400Hz ～ 2000Hz、D は 4000Hz ～ 5600Hz の特性である。この放射特性図より明らかなように、人間が発生する音声は、話者正面方向に最も強く放射され、上方や下方、及び左右方向に放射される音声のパワーは、話者正面方向に比べて小さい。

それゆえ、図 1 のように第 1 のマイクロホン 1 1 を話者の顔の真上天井に配置し、第 2 のマイクロホン 1 2 を第 1 のマイクロホン位置より 1 ～ 5 cm 程度後頭部側の天井に配置すれば、① 2 つのマイクロホン 1 1、1 2 で受音するノイズのパ

ワーをほぼ同一にできる一方、②2つのマイクロホン11, 12で受音する音声パワーを異ならせることができる。すなわち、2つのマイクロホン11, 12の出力信号に含まれるノイズ $Xn_1(z)$, $Xn_2(z)$ を略等しくできると共に、2つのマイクロホン11, 12の出力信号に含まれる音声信号 $Xs_1(z)$, $Xs_2(z)$ を異ならせることができ、 $[Xn_1(z)/Xn_2(z)] \neq [Xs_1(z)/Xs_2(z)]$ とすることが可能である。

【0013】

(c) 動作

マイクロホン11, 12にノイズのみが入力する非音声認識時において、適応信号処理部3は適応信号処理により次式

$$En(z) = Xn_1(z)z^{-d} - Xn_2(z)W(z) \quad (1)$$

において、 $\{En(z)\}^2$ の平均値が最小となるように適応フィルタ3bのフィルタ係数 W を決定する。

一方、音声認識時、適応信号処理部3はフィルタ係数の更新をせず、前記非音声認識時に決定したフィルタ係数 W を適応フィルタ3bに設定して音声信号を出力する。この場合、マイクロホン11, 12の出力に含まれる音声信号 $Xs_1(z)$,

$Xs_2(z)$ は異なり、 $[Xn_1(z)/Xn_2(z)] \neq [Xs_1(z)/Xs_2(z)]$ となるため、次式

$$Es(z) = Xs_1(z)z^{-d} - Xs_2(z)W(z) \quad (2)$$

により求まる音声出力 $Es(z)$ は最小値には(ノイズと異なり、あまり小さくは)ならない。

以上より、(1)式のノイズ出力 $En(z)$ のパワーが0となるように適応フィルタ係数 W を決定しても、(2)式の音声出力 $Es(z)$ はノイズと同様に小さくはならず、音声信号のSN比を改善することができる。

【0014】

以上要約すれば、図1に示すように、マイクロホン11, 12を比較的近距離に配置することで2つのマイクロホンが出力するノイズ間のコヒーレンスの低下を少なくし、更に、図2に示すような人間の音声放射特性を考慮することで、マイクロホン11, 12が比較的近距離に配置されているにも拘らず、1つのマイクロホン11はできるだけ高いSN比で音声拾い、もう一方のマイクロホン12では、できるだけ低いSN比で音声拾う。この結果、ノイズ出力が零となるよう

に適応フィルタ係数 W を決定しても、音声出力はノイズと同様に小さくはならず、音声信号のSN比を改善することができる。

【0 0 1 5】

(d) マイクロホン位置とSN比改善量の検討

図2の放射特性より人間が発声し空間に放射した音声は、後頭部側で特に大きく減衰し、正面に放射される音声に比べてレベルが小さくなることがわかる。それゆえ、本発明のマイクロホンシステムは、図1に示したように人間の頭部真上付近から後頭部側にかけて設置するのが基本であり、このように第1、第2のマイクロホン11、12を設置することによりSN比をより大きく改善することができる。

図3はペアマイクロホン位置の説明図、図4は図3の各ペアマイクロホン位置におけるSN比改善量を示す図である。図3に示すように、3 cm間隔のペアマイクロホン11、12を複数の位置①、②、③に設置し、それぞれのSN比の改善量(どの程度SN比が向上するか)を、1500ccの乗用車(セダン)で調べてみると、図4に示す結果が得られた。この図4より、ペアマイクロホン11、12を①の位置に設置したとき、すなわち、「1つのマイクロホンを話者10の顔のほぼ真上に設置し、もう1つのマイクロホンを少し離して後頭部側に配置したとき」にSN比の改善量が最も高くなることがわかる。

【0 0 1 6】

図5はペアマイクロホン間隔の説明図、図6は図5の各ペアマイクロホン間隔におけるSN比改善量を示す図である。図5に示すように、第1のマイクロホン11を話者10の顔のほぼ真上に固定し、第2のマイクロホン12を後頭部側にそれぞれ3 cm, 6 cm, 9 cm, 12 cm離して配置し、最適なマイクロホンの間隔について調べてみると、図6に示す結果が得られた。この図6より、2つのマイクロホン11、12の間隔が狭いほど、SN比の改善量が高いと考えられる。しかし、図1に示したシステムでは、間隔を0 cmとするとノイズを完全に消去できるが、音声もまた完全に消去してしまう。このため、音声受音システムとして機能しないことになる。また、小型マイクロホンといえどもそれ自体の大きさがあるため、マイクロホンどうしを完全にくっつけても、マイクロホンの中心間隔は1cm

程度より小さくはならない。それゆえ、マイクロホンの間隔は、車種の違いとマイクロホンの大きさにより若干の幅があるものの、せいぜい1～5cm程度にするのがよい。

【0017】

図7は発声者別のSN比改善量説明図である。この図7より明らかなように、本発明のマイクロホンシステムでは、人の違いによる性能(SN比改善量)のばらつきは1dB程度であり、話者の違いによる影響は少ない。

以上2つのマイクロホンを話者の頭上に配置した場合について説明したが、2つのマイクロホンを比較的近距離に配置し、1つのマイクロホンにより、できるだけ高いSN比で音声を拾い、もう一方のマイクロホンにより、できるだけ低いSN比で音声を拾うようにできれば、配置位置は頭上に限らない。

以上、本発明を実施例により説明したが、本発明は請求の範囲に記載した本発明の主旨に従い種々の変形が可能であり、本発明はこれらを排除するものではない。

【0018】

【発明の効果】

以上本発明によれば、各マイクロホンを接近して配置すると共に、一方のマイクロホンから出力する信号のSN比を高くし、他方のマイクロホンから出力する信号のSN比を低くするようにしたから、ノイズ出力が最小となるように適応フィルタ係数を決定しても、音声出力は零とならず、音声信号のSN比を改善することができる。すなわち、少ないマイクロホン数であるにも拘らず、高いSN比で音声を受音出力することができる。

又、本発明によれば、一方のマイクロホンを話者の顔の真上天井に配置し、他方のマイクロホンを該一方のマイクロホンの位置より1～5cm程度後頭部側に離して配置することにより、マイクロホンが比較的近距離に配置されているにも拘らず、1つのマイクロホンはできるだけ高いSN比で音声を拾い、もう一方のマイクロホンでは、できるだけ低いSN比で音声を拾うようにできる。

【図面の簡単な説明】

【図1】

本発明のマイクロホンシステムの構成図である。

【図 2】

人間の音声放射特性図である。

【図 3】

ペアマイクロホンの位置説明図である。

【図 4】

ペアマイクロホン位置とSN比改善量の関係図である。

【図 5】

ペアマイクロホンの間隔説明図である。

【図 6】

ペアマイクロホンの間隔とSN比改善量の関係図である。

【図 7】

発声者別SN比改善量説明図である。

【図 8】

SN比と認識率の関係図である。

【図 9】

従来のマイクロホンを 2 つ使用した場合の高SN比受信システムである。

【図 1 0】

目標応答設定部の特性図である。

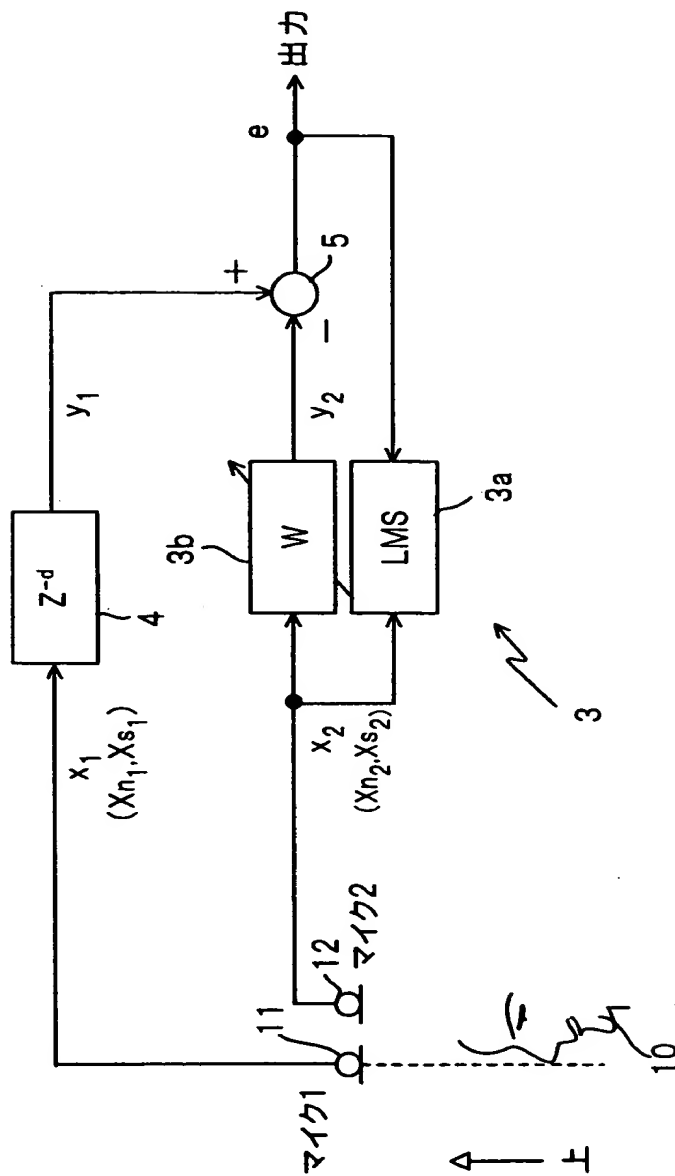
【符号の説明】

- 1 1, 1 2・・・第 1、第 2 のマイクロホン
- 3・・・適応信号処理部
- 3 a・・・LMS演算部
- 3 b・・・適応フィルタ
- 4・・・目標応答設定部
- 5・・・減算部

【書類名】 図面

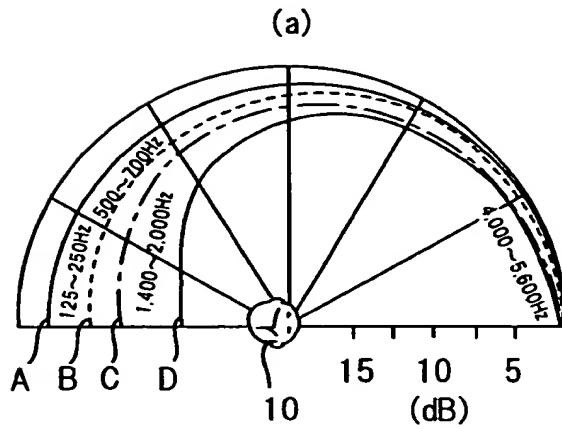
【図 1】

本発明のマイクロホンシステムの構成

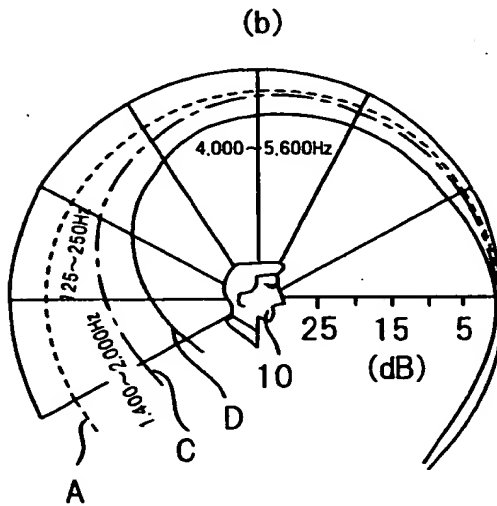


【図2】

人間の音声放射特性

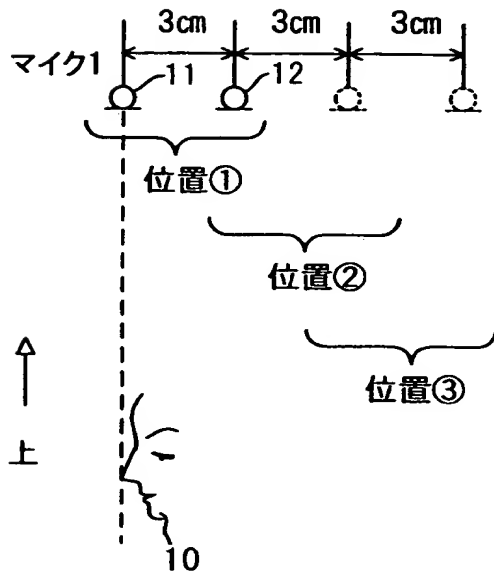


A=125~250Hz
B=500~700Hz
C=1,400~2,000Hz
D=4,000~5,600Hz



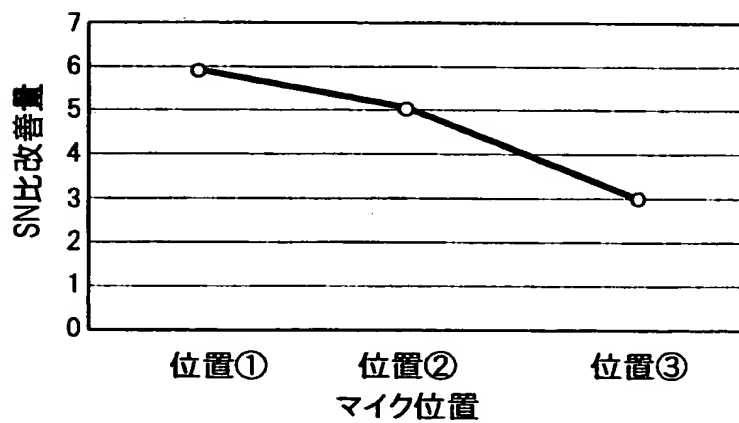
【図3】

ペアマイクロホンの位置説明図



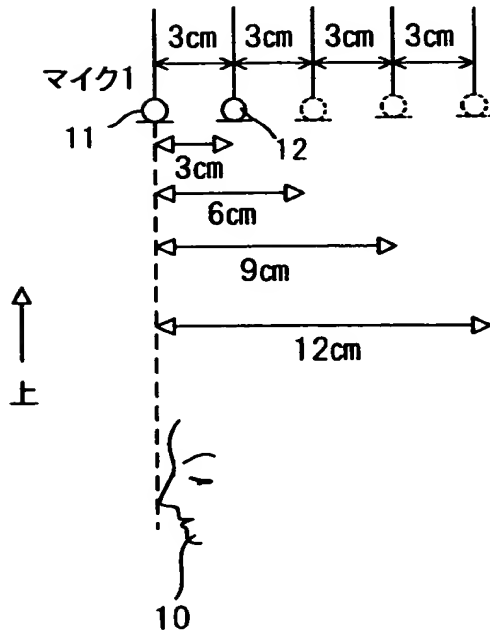
【図4】

ペアマイクロホン位置とSN比改善量の関係図



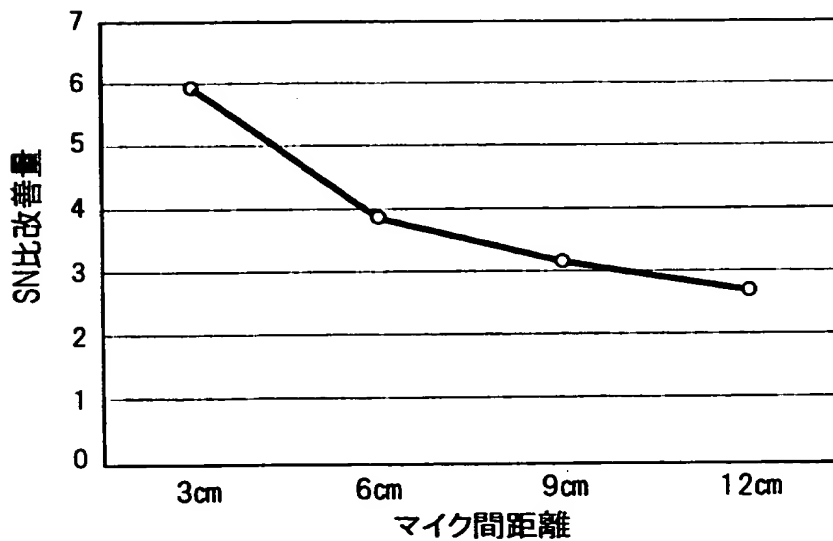
【図 5】

ペアマイクロホンの間隔説明図



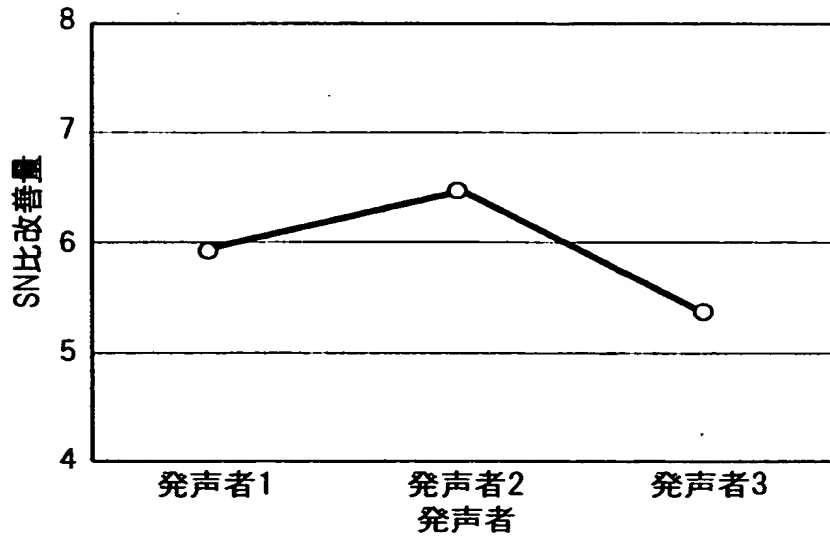
【図 6】

ペアマイクロホンの間隔とSN比改善量の関係図



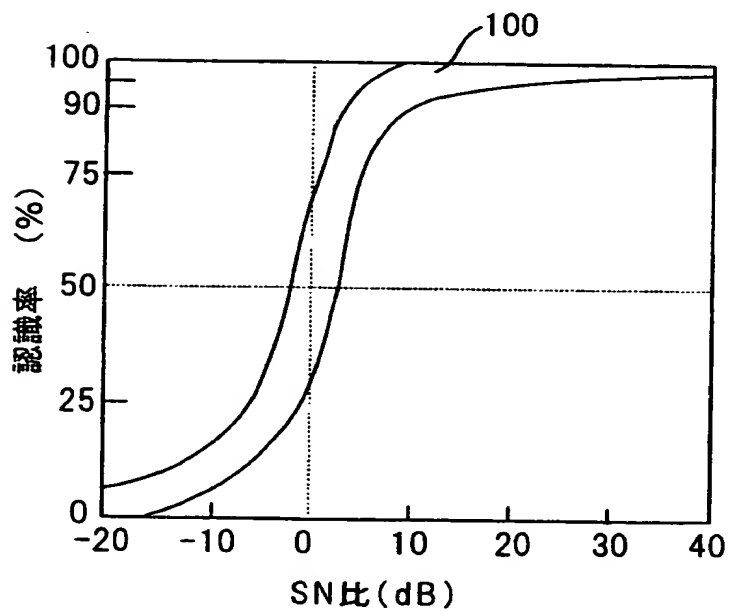
【図7】

発声者別SN比改善量



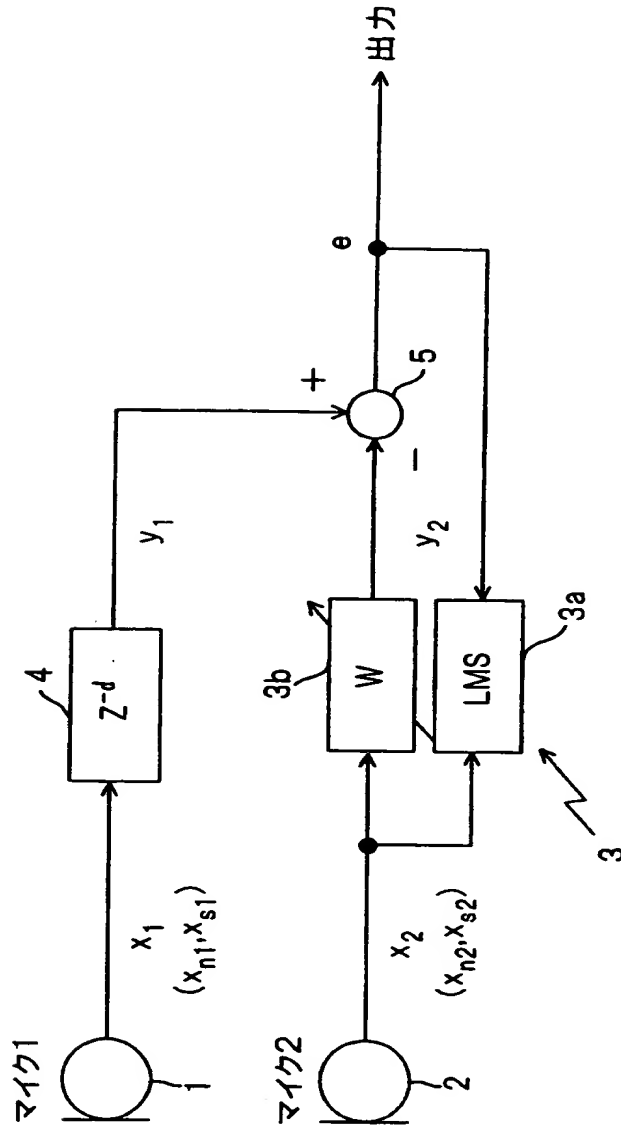
【図8】

SN比と認識率の関係



【図9】

従来のマイクロホンを用いた場合の高SN比受信システム

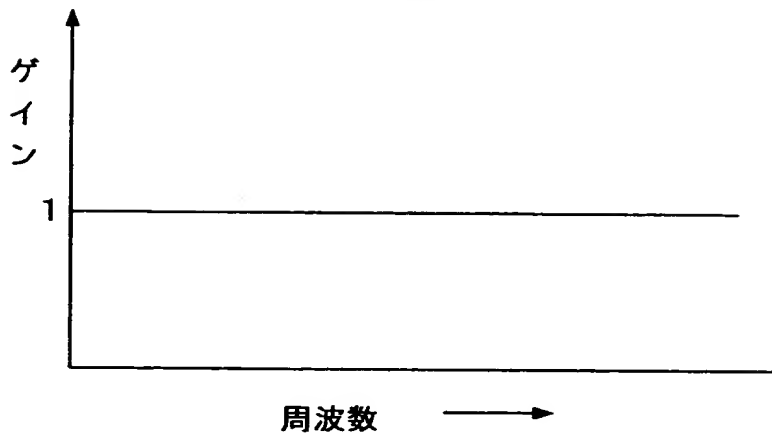


【図 10】

目標応答設定部の特性

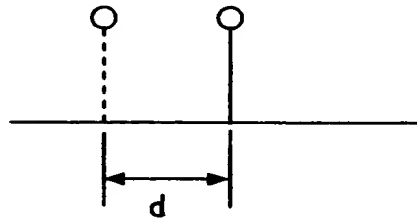
(a)

周波数特性



(b)

インパルス応答特性



【書類名】 要約書

【要約】

【課題】 マイクロホン 2 つ使用するマイクロホンシステムにおいて、音声信号の SN 比を改善する。

【解決手段】 2 つマイクロホン 11, 12 の出力信号を用いて適応信号処理を行い SN 比を改善した話者音声信号を出力するマイクロホンシステムにおいて、各マイクロホン 11, 12 を接近して配置すると共に、一方のマイクロホンから出力する信号の SN 比を高くし、他方のマイクロホンから出力する信号の SN 比を低くする。例えば、一方のマイクロホン 11 を話者 10 の顔の真上に配置し、他方のマイクロホン 12 を該一方のマイクロホンの位置より 1 ～ 5 cm 程度後頭部側に離して配置する。

【選択図】 図 1

認定・付加情報

特許出願の番号	平成11年 特許願 第121518号
受付番号	59900412261
書類名	特許願
担当官	茨田 幸雄 6051
作成日	平成11年 5月14日

<認定情報・付加情報>

【特許出願人】

【識別番号】	000101732
【住所又は居所】	東京都品川区西五反田1丁目1番8号
【氏名又は名称】	アルパイン株式会社
【代理人】	申請人
【識別番号】	100084711
【住所又は居所】	千葉県千葉市花見川区幕張本郷1丁目14番10号 幸栄パレス202 齋藤特許事務所
【氏名又は名称】	齋藤 千幹

出 願 人 履 歴 情 報

識別番号 [000101732]

1. 変更年月日 1990年 8月27日

[変更理由] 新規登録

住 所 東京都品川区西五反田1丁目1番8号

氏 名 アルパイン株式会社